Aus dem Institut für Medizinische Psychologie
der Universität Düsseldorf
Direktor: Prof. Dr. H.-J. Steingrüber

DISCRIMINATION OF NEONATE CRIES BY MOTHERS, NON-MOTHERS AND COMPUTER ANALYSIS

Dissertation
zur Erlangung des Grades eines Doktors der Zahnmedizin
der Medizinischen Fakultät der Universität Düsseldorf

vorgelegt von
Gabriele Nolten

1983

UNIV.
LIBRARY

Aus dem Institut für Medizinische Psychologie der
Universität Düsseldorf
Direktor: Prof. Dr. H.-J. Steingrüber

DISCRIMINATION OF NEONATE CRIES BY MOTHERS, NON-MOTHERS AND COMPUTER ANALYSIS

Dissertation

zur Erlangung des Grades eines Doktors der Zahnmedizin
der Medizinischen Fakultät der Universität Düsseldorf

vorgelegt von

Gabriele Nolten

1983

Gedruckt mit Genehmigung der Medizinischen Fakultät
der Universität Düsseldorf

Der Dekan
der Medizinischen Fakultät
der Universität Düsseldorf
gez. :
Prof. Dr. A. Hopf

Referent
gez.:
Prof. Dr. H.-J. Steingrüber

Korreferent
gez.: Prof. Dr. H. Böttger

Discrimination of Neonate Cries By Mothers, Non-Mothers and Computer Analysis

Gabriele Nolten, Rudolf Bisping,
and Hans-Joachim Steingrueber
Institut fuer Medizinische Psychologie
Universitaet Duesseldorf

Summary

Nine mothers had to judge the cries of neonates on the third, fourth and fifth day after delivery in a complete comparison situation. Each mother had to compare the cry of her own infant (anchor cry) with the cries of four other newborn babies of the same age. The discrimination abilities of the mothers were set against two reference judgements by (1) non-mothers, who were familiarized with the anchor cries and (2) computer analysis, based on mean fundamental frequency and harmonic structure. The results show that the discrimination ability of the mothers improves considerably from the third to the fifth day after delivery when it reaches the optimum level defined by computer analysis. Performance by non-mothers is equally high under all conditions but does not achieve discrimination by computer analysis or the performance by mothers on the fifth day after delivery. Since performance by mothers and non-mothers is clearly distinct in respect to correlation with computer analysis scaling, it is justified to conclude that the two groups of subjects base their judgements on different cry characteristics.

Summary

Introduction

The identification of own babies by their cries necessitates (1) the capability of discriminating between different cries on the part of the mother and (2) the existence of an individual cry pattern largely invariant in different situations. Morsbach and Murphy (1979) studied different groups of subjects (midwives, parents with children below and above 1 year and non-parents who sometimes have contact or have no contact with babies or small children), who had to discriminate the cries of 2 to 5 day old newborn babies. The results showed that the percentage of correct discrimination by midwives (57%) was considerably higher than that of the other four groups for which no significant differences were found (correct discrimination approx. 30%). The authors explained the difference between the groups by arguing that irrespective of individual experience in handling young children, all persons have certain abilities of discriminating between acoustic stimuli (in this case cries of unknown infants) and only considerable experience (midwives) would improve these abilities. Work has gone beyond the recognition of infants by their cries and investigated the discrimination of certain emotional states expressed by different cries. Results in this area are, however,

still ambiguous (see Murray, 1979). Regarding the second condition, the existence of individualistic invariable cry patterns, Truby and Lind (1965) have shown, for instance, for 17 cries of six infants that each baby can be recognized as reliably by his characteristic cry pattern as by his fingerprints justifying, in fact, the use of the term of cry prints. Work on neonates by Ostwald (1963) and Ringel and Kluppel (1964) came to the same conclusions. These researchers found a very high intra-individual invariance of certain cry characteristics and significant inter-individual differences. Data on the frequency of the correct identification of the newborn babies by their own mothers during the first week range from 31% (Valanne, Vuorenkoski, Partanen, Lind, & Wasz-Hoeckert, 1967) to approx. 80% (Morsbach, 1980; Morsbach & Bunting, 1979) and even to 100% (Formby, 1967); for 3- to 5-month old infants, the identification ability averages 91% according to Murry, Hollien, and Mueller (1975). In these studies mothers were asked to listen to a randomized sequence of cries of their own babies and cries of other babies on a tape and to identify the cries of their own babies. The researchers examined how many mothers were significantly successful in identifying their own children.

It is difficult to explain the considerable scatter of these results. On the one hand, the method is rather inaccurate, if compared with similarity scaling; on the other hand, it is impossible to evaluate, for instance, the effect of the different number of yes/no decisions (total number of cries or number of tests) or the largely uncontrolled quality of the cry stimuli on the results. Finally, the work of Formby (1967) has shown that during the first few days after delivery, the duration of the mothers exposure to her baby may be significant: 100% identification was only achieved by mothers between the third and the sixth day of the baby's life, while the identification rate was only approx. 50% before the third day. Research published is therefore not yet sufficient to identify the conditions which are instrumental for the mothers to be able to recognize her own baby by his voice. Our investigations were therefore designed to answer the following questions:

1. What is the accuracy of identification of newborn babies by their own mothers on the third, fourth and fifth day after delivery?

2. How accurate is the discrimination performance by mothers relative to trained non-mothers and the evaluation of certain physical cry characteristics by spectral computer techniques?

3. What is the correlation between the discrimination abilities of mothers and non-mothers and invariant cry characteristics?

Satisfactory answers to these questions can only be given, if (1) the response of the subjects (mothers/non-mothers) as well as (2) specific invariant cry characteristics can be recorded reliably on a graded scale. Among the invariant characteristics important for the recognition of voice signals, we selected pitch (fundamental frequency) and timbre (relative number of harmonics).

Method

Cry stimulus material

The cries used were the cries of 9 (5 male and 4 female) newborn babies of the mothers who volunteered for the test (see 'Judgements by Mothers') and who had to identify their own infants. The infants were healthy and their weights were between 2000 and 4200 grammes. The cries of these test infants were set against the cries of 12 other children. All cries were recorded during the morning bath in the nursery under standardized conditions (Uher CR 210 cassette tape recorder, Uher M 543 microphone, distance from

microphone approx. 40 cm). The method used was the complete comparison technique and each test consisted of the cry of the mother's own baby (anchor cry) and four cries of other infants (reference cries). The

$$\frac{N\,(N-1)}{2} = 10$$

N = 5 cries were dubbed onto a test tape with pairs. Each of the cries was hence heard four times. To monitor the effect of the sequence of the cries, each cry was twice recorded in second place. A number was assigned to each of the N = 10 pairs and the sequence was permuted in the order of 1, 2, N, 3, N-1, 4, N-2, etc. . On the tape, each of the pairs was announced by a number (1 to 10). The pairs were separated from each other by a short sine tone (440 Hz). The three tests consisted of cries recorded on the third, fourth and fifth day after delivery. If, for instance, the cry of one of the newborn babies of the subject mothers (anchor cry) was recorded on the third day after delivery, this cry was mixed on the tape with the reference cries of four other babies of the same day. The anchor cries were again recorded on the fourth and on the fifth day after delivery and then again paired with the cries of 4- respectively 5-day old babies. The babies whose cries were paired with the anchor cries recorded on the third, fourth and fifth day were different so that a total of 3 x 4 = 12 other babies were included in the test. The tapes were always

prepared for the anchor cries recorded in the morning to be added at specific points on the tapes.

All subjects (mothers and non-mothers) heard the tape through earphones at a maximum sound pressure level of 85 dB(A). To eliminate atypical spectral characteristics of the 12 reference cries, they were selected by means of a homogeneity criterion from 21 cry samples. Most of these cry samples had a duration of several minutes. A continuous passage of 15 seconds containing the sound pressure peaks of the sample was used for the selection. The signals were read into a laboratory computer (CP/M-compatible Z-80 ECB-Bus system) by means of a 12 bit analog-digital conversion with a sample rate of 10 kHz. The average power spectrum was determined for the 15-second-recording of each of the 21 cries by means of an FFT programme (see Bergland & Dolan, 1978), supported by arithmetic processing hardware. The analysis was made for the 0-5 kHz frequency range. The resolution was 512 frequency points giving an approx. 9.8 Hz/frequency point bandwith. The 21 spectra were again averaged and the average deviation of each spectrum from the average spectrum was determined. The 12 reference cries used were cries whose spectral deviation from the average spectrum was in the normal range of one standard deviation.

Judgements by mothers

The subjects were 9 secundiparous mothers with uncomplicated pregnancies and deliveries (age range: 22 to 36 years). These mothers contacted their infants during meals five times per day for approximately half an hour. The method of judgement was the complete paired comparison technique. The instruction given to the mothers was as follows: "Please, judge which one of the two cries is more similar to the cry of your baby!" Each mother did three tests consisting of ten pairs each. These tests were run in the evening of the third, fourth and fifth day after delivery.

Judgements by non-mothers

Judgements by the non-mothers had to be comparable with the judgement of each mother (mothers 1-9) during each test on days 1 to 3 (day 3 to 5 after delivery), thus providing a standardization criterion. The subjects were 8 female students without infants of their own (age range: 22 to 26 years). The cry samples used for the reference group tests were the same cry pairs as the cry stimuli to be judged by the mothers. Each non-mother had to analyse 9 (mothers) x 3 (days after delivery) x 10 (tests) = 270 pairs.

Table 1

Judgements to be Replicated by Each Non-Mother:

Grouping of the Test Sessions.

	Mother								
	1	2	3	4	5	6	7	8	9
1. Test day	10	10	10	10	10	10	10	10	10
2. Test day	10	10	10	10	10	10	10	10	10
3. Test day	10	10	10	10	10	10	10	10	10
Number of judgements per test session	30	30	30	30	30	30	30	30	30
							Total:		270

The 270 individual judgements to be made were grouped in 9 test sessions consisting of 3 x 10 tests = 30 pairs each (see Table 1). The anchor cries to be identified were the same 9 (mothers) x 3 (days after delivery) = 27 cries to be identified by the mothers. As the non-mothers were not familiar with these cries, they were exposed to the anchor cries in an approach designed to maximize familiarization. During the training, the non-mothers first had to learn three anchor cries. These cries were always cries by three different children recorded on three different days

(third, fourth and fifth day after birth). During the first training session, the non-mothers heard the three anchor cries through earphones. During the next session, the sequence of these cries was changed and the non-mothers were asked to note the sequence on an answer sheet within 10 seconds. They were then informed about the correct sequence. Familiarization was considered complete, when the correct sequence had been identified during 4 training sessions in succession. Following familiarization, the non-mothers three times heard one of the cries they had been familiarized with. They were told that this cry was to be scaled in a way analogous to the scaling procedure of the mothers during the subsequent test consisting of 10 pairs. The non-mothers then heard the test tape that had been listened to by the mother on one of the three test days. When the non-mother had made the 10 judgements, she heard another one of the three cries she had been familiarized with and then listened to another test tape. The same procedure was used for the third cry to be identified. The nine familiarization sessions were distributed over five days. They consisted of two training sessions on the first four days and one training session on the fifth day. Systematic permutation was used to eliminate any sequential effects. To eliminate the effect of fatigue, two

familiarization sessions were separated by a break of 20 minutes.

Comparisons by computer analysis

To compare the judgements made by the mothers and the non-mothers with an objective scale, a paired comparison based on the physical characteristics of the cries was made for the N = 39 cries (12 reference cries, 27 anchor cries). The physical characteristics used for this test were the mean fundamental frequency and the harmonics of the cries. The numerical base for paired comparison was (1) the absolute frequency difference between two cries in terms of fundamental frequency and (2) an acoustic distance measure in the case of the harmonics, using the equation $D = \left(\sum_{i=1}^{m} (A_i - R_i)^2 \right)^{1/2}$ where A_i and R_i represent the amplitudes of the two spectral lines of the i-th harmonic to be compared. Nine was used for m, since the first nine harmonics are, according to our experience, sufficient to represent the spectral structure of a baby cry in the frequency range analyzed.

Results

The BTL scaling technique (Bradley & Terry, 1952) was used for comparing mother/non-mother judgements and computer analysis. The variance due to the linear

(transitive) component was estimated and tested for significance by a χ^2-procedure (Mosteller, 1951). The estimate indicates in percent to what degree the choice depends upon just one dimension of the choice behaviour. This analysis revealed that in all cases of BTL-scaling (Figs. 1 to 4), the linear variance component was significant (p .001), accounting for about 80% to 90% of the total variance of the data. To determine the consistency of the judgements of the individual mothers, the Kendall coefficients of consistency (Kendall, 1955) were calculated. The coefficients are summarized in Table 2.

Table 2

Consistency Coefficients for the Judgement by the Mothers for Test Day 1 to 3

Mother	1. Test day	2. Test day	3. Test day
1	1.0	1.0	1.0
2	1.0	1.0	1.0
3	1.0	.6	1.0
4	1.0	.6	1.0
5	1.0	1.0	1.0
6	1.0	.8	.8
7	.2	.8	.8
8	.6	1.0	1.0
9	1.0	.8	1.0

Most coefficients are reasonably satisfactory and fairly good reliability may be assumed. Some coefficients are below an acceptable level of consistency, as, in particular, in the case of mother no. 7. For the non-mothers the coefficients of consistency were quite comparable to those of the mothers, ranging from .6 to 1.0. Table 3 shows the identificaton rates for the mothers and the non-mothers.

Table 3

Identification Rates for Mothers and Non-Mothers for Test Day 1 to 3.

	1. Test day	2. Test day	3. Test day
Mothers	81%	89%	97%
Non-mothers	97%	95%	94%

Due to the scaling method used, the identification rates could be calculated simply by comparing the number of correct judgements with the total number of judgements possible. The number of correct judgements by non-mothers ranged between 94% and 97%. Judgements by mothers improved distinctly as a function of time,

but even on the first test day, three days after delivery, the identification rate of 81% was fairly high. As a trivial result -not included in the table- the computer analysis allowed an identification rate of 100% since the information given by this kind of analysis is very specific. Figs. 1 to 4 show the similarity scaling data as a function of the five cry stimuli on the three test days. The lowest BTL-value on each test day was the zero similarity reference cry (since the BTL-scale is an interval scale, it is not possible to show the distance from an absolute reference point).

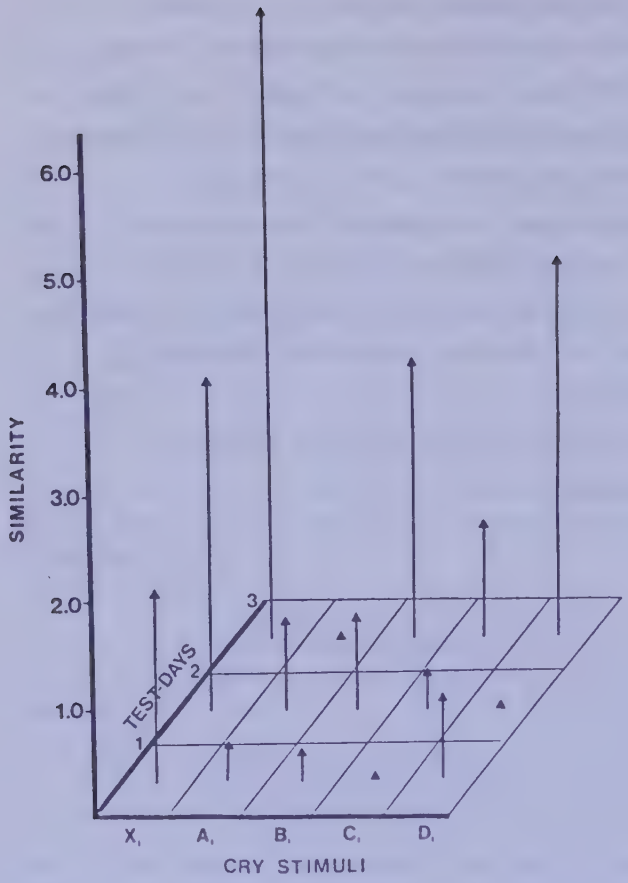

Fig. 1: Similarity scaling (BTL-scale) of mothers as a function of various cry stimuli and test days (X: anchor cry; A_i - D_i: reference cries).

Fig. 1 shows that the similarity scaling for mothers concerning the cries of their own babies steeply rises

as a function of time, peaking at a maximum BTL-value of about 6.3 on the third day.

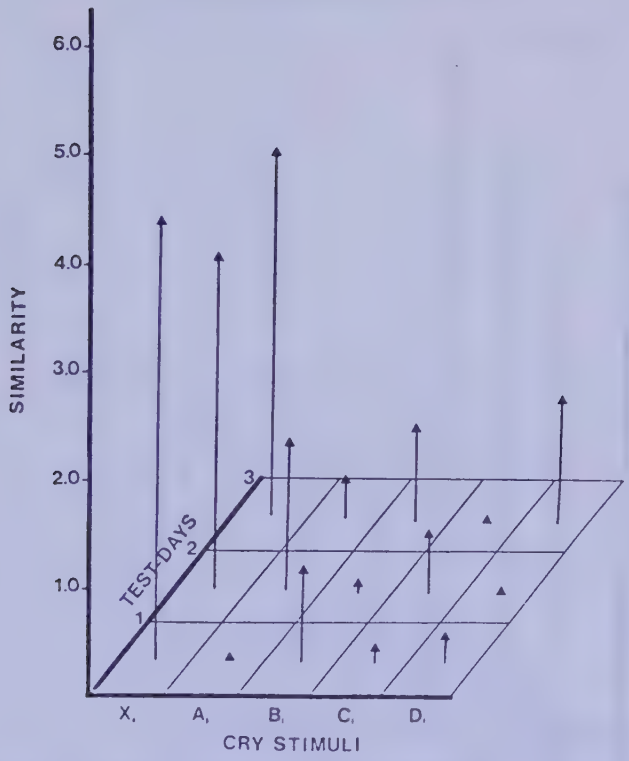

Fig. 2: Similarity scaling (BTL-scale) of non-mothers as a function of various cry stimuli and test days (X_i: anchor cry; A_i - D_i: reference cries).

The similarity scaling for non-mothers is fairly high, but does not vary as a function of time (see Fig. 2). Figs. 3 and 4 show the similarity scaling based on computer analysis as a function of harmonics and fundamental frequency.

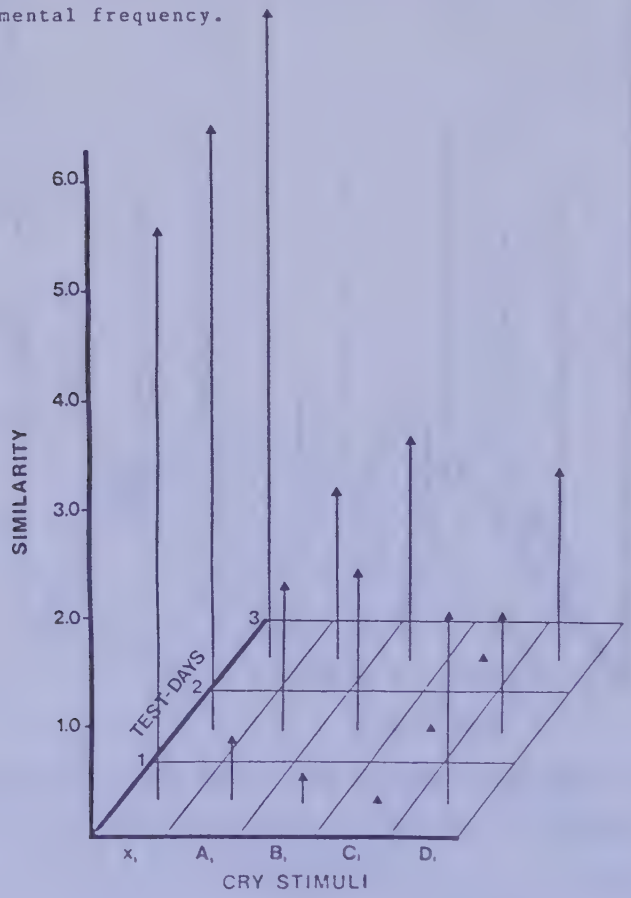

Fig. 3: Similarity scaling (BTL-scale) of harmonics as a function of various cry stimuli and test days (X_i: anchor cry; A_i - D_i: reference cries).

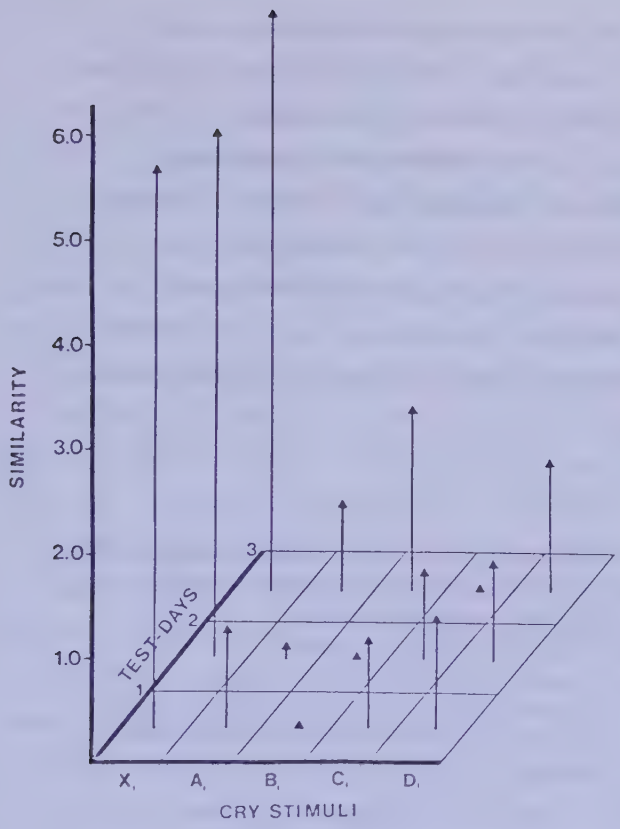

Fig. 4: Similarity scaling (BTL-scale) of fundamental frequency as a function of various cry stimuli and test days (X_i: anchor cry; A_i - D_i: reference cries).

Since the identification of the anchor cries is 100 % with both variables and no inconsistent values

appeared in the paired comparison matrices, the distances of the anchor cries from the reference cries may be considered as maximum discrimination values. For an insight into the statistical relationship between subjective judgements as dependant variables (discrimination by mothers and non-mothers) and objective judgements based upon harmonics and fundamental frequency as predictor variables, multiple correlations between these two sets of judgements were calculated.

Table 4

Multiple Correlations of Computer Analysis Data With Judgements by Mothers and Non-Mothers.

	1. Test day	2. Test day	3. Test day
Mothers	.53	.46	.77
Non-mothers	.76	.68	.90

Table 4 shows that the correlations between the judgements by the non-mothers and the combined physical characteristics are somewhat higher than the same correlations fo the judgements by the mothers. Obviously, the correlations are higher for the third day than for the second day and the first day.

Discussion

Although the mothers' identification rate of 81% on the third day after delivery is very high, identification continued to improve until the fifth day. This result is somewhat different from earlier results regarding the correlation between identification ability and the age of the infant. In Formby's study (1967) there was no improvement after the second day, since all subject mothers identified their children correctly between the third and the sixth day. In other studies published (Morsbach, 1980; Morsbach & Bunting, 1979; Valanne et al., 1967), no correlation seems to exist between identification capability and the age of the babies. Naturally the data used in our study and the data reported by other researchers are not directly comparable. In our research, the same sample of judges was tested on three different days and the average improvement in identification ability was determined. In all other cases, different samples were taken on the different days after birth. If it is assumed that the identification abilities of different mothers differ more widely than the identification rates on successive days the different results obtained by various authors could, in fact, be explained. Since the samples of

mothers for each day after birth are very small, there is a considerable risk that progress made by mothers in the identification of their own babies cannot be recognized because of the different identification abilities of different mothers. Since the sample remained the same in the case of our research and the same mothers went through the same test on three successive days, the error due to different identification abilities of different samples of mothers was eliminated and progress made by individual mothers in the identification of their own children could be shown more clearly. These considerations lead to the unanswered question, to what extent mothers are really capable of discriminating between different cries and whether this ability is different from that of non-mothers. Since correct identification neccessitates correct discrimination the point in question is not only the correct identification of specific cries, but also the correct placement of all cries on a similarity scale. Work so far published on the identification of neonates by their own mothers during the first week of the baby's life has always been based on the significance of identification rates. The quantitative approach used in this paper goes beyond the merely qualitative approach and may therefore help to clarify the reasons for different

identification abilities measured: The discrimination ability can be represented on an interval scale and the reliability (consistency) of individual and group judgements can be assessed. Moreover the correlation between scaled judgements and physical cry characteristics can be determined.

Our results show that the mean consistency of judgements is quite satisfying, i.e. the capability of discriminating, and both groups do not differ substantially from each other. This result somewhat modifies the conclusions drawn by Morsbach and Murphy (1979): Apparently, the essential criterion is not the experience of parents and non-parents with regard to the amount of contact with children, but familiarity with a cry on which they base their judgements. At the same time, the investigations show that both in the group of mothers and in the group of non-mothers, there were persons with good and persons with less good discrimination abilities. The consistency coefficients for mothers were between .2 and 1.0, while they ranged from .6 to 1.0 for non-mothers. If the change in performance during the three days is considered, the data show that performance by mothers improves considerably, while performance by non-mothers remained largely unchanged. This result can be explained easily for the non-mothers: The permutation of the stimuli

prevented the sequence of stimuli from being the same as for the group of mothers. If identification or discrimination by non-mothers had also differed between the three test days, the difficulties of identifying the cries would not have been the same on the three days for either group. Since this was not the case, the improvement in performance by mothers during the three test days is due to improved discrimination abilities. The mothers' discriminaton performance on the third day will be the best performance that can be achieved, if discrimination by computer analysis is taken as a yardstick. The multiple correlations indicate that there is a close correspondance between the linearly combined similarity scaling based on harmonics and fundamental frequency and the discrimination by mothers, respectively non-mothers. The harmonics of an acoustic signal correspond to the timbre of the voice heard, while fundamental frequency corresponds to the pitch in the voice. It is exactly because both physical properties of an acoustic signal are essential components of the cry pattern they were used for the computer analysis. Comparing the different sets of correlations wih each other it can be seen, that the discrimination by non-mothers is correlated to a higher degree with the computer analysis than the discrimination by mothers. Obviously the judgements of

the mothers and non-mothers are based on different cues of the cry signal. The use of these two physical parameters in the present study is to be considered as a first approach to the general question on which signal characteristics the discrimination performance is based on. To answer this question a broader spectrum of variables, such as time changes in harmonic composition, intonation patterns, analysis of voiced vs. unvoiced components etc. , are to be taken into account. This kind of analysis would have gone beyond the scope of the present paper and is rendered to a report in preparation.

References

Bergland, G. D. , & Dolan, M. T. (1979). Fast Fourier algorithms. In Digital Signal Processing Committee (Ed.), Programs for Digital Signal Processing (pp. 1.2-1 - 1.2-18). New York: IEEE Press.

Bradley, R. A. & Terry, M. E. (1952). The rank analysis of incomplete block designs I: The method of paired comparisons. Biometrica, 39, 324-345.

Formby, D. (1967). Maternal recognition of infants' cry. Developmental Medicine and Child Neurology. 9, 293-298.

Kendall, M. G. (1955). Further contributions to the theory of paired comparisons. Biometrics, 11, 43-62.

Morsbach, G. (1980). Maternal recognition of neonates' cries in Japan. Psychologia, 23, 63-69.

Morsbach, G. , & Bunting, C. (1979). Maternal recognition of their neonates'cries. Developmental Medicine and Child Neurology, 21, 178-185.

Morsbach, G. , & Murphy, M. C. (1979). Recognition of individual neonates'cries by experienced and inexperienced adults. Journal of Child Language, 6, 175-179.

Mosteller, F. (1951). Remarks on the method of paired comparisons III: A test of significance for paired comparisons when unequal standard deviations and

equal correlations are assumed. Psychometrika, 16, 207-218.

Murray, A. D. (1979). Infant crying as an elicitor of parental behavior: An examination of two models. Psychological Bulletin, 86, 191-215.

Murry, T. , Hollien, H. , & Mueller, E. (1974). Perceptual responses to infant crying: Maternal recognition and sex judgments. Journal of Child Language, 1, 89-95.

Ostwald, P. (1963). Soundmaking: The acoustic communication of emotion. Springfield, IL. : Charles C. Thomas.

Ringel, R. , & Kluppel, D. (1964). Neonatal crying: A normative study. Folia Phoniatrica, 16, 1-9.

Truby, H. M. , & Lind, J. (1965). Cry sounds of the newborn infant. Acta Paediatrica Scandinavica, 163, 7-59.

Valanne, E. H. , Vuorenkoski, V. , Partanen, T. J. , Lind, J. , & Wasz-Hoeckert, O. (1967). The ability of human mothers to identify the hunger cry signals of their own newborn infants during the lying-in period. Experientia, 23, 768-769.

Herrn Prof. Dr. H.-J. Steingrüber danke ich herzlich für die vielen Anregungen, Ermunterungen und hilfreichen Diskussionen bei der Vorbereitung und Durchführung der Arbeit.
Herrn Dr. R. Bisping danke ich für zahlreiche Hinweise und Hilfen vor allem bei der Lösung methodischer und technischer Probleme.

Lebenslauf

Name:	Gabriele Nolten
23. 8. 1956	geboren in Heinsberg als Tochter der Eheleute Theo und Gertrud Nolten
1963 - 1966	Besuch der Grundschule (St. Georg - Schule) in Wassenberg
1966 - 1975	Besuch des Städtischen Mädchengymnasiums in Erkelenz
1975	Abitur
1975 - 1980	Studium der Zahnmedizin an der Universität Düsseldorf
1980	Staatsexamen
1981 -	Wiss. Angestellte an der Abteilung für Zahnärztliche Prothetik und Defektprothetik der Westdeutschen Kieferklinik der Universität Düsseldorf (Direktor: Prof. Dr. H. Böttger).

FOTODRUCK J. MAINZ GmbH
Neupforte 13, Nähe Markt, D-5100 Aachen
Telefon: (0241) 27305 und (0241) 38606